

Title of the course: Computational Multimodal Analyses and Archiving

Name of the teacher: Wim Pouw and Babajide Owoyele

Contact details of the teacher

Affiliation: Donders Institute for Brain, Cognition and Behaviour, Nijmegen /Hasso Plattner Institute Potsdam

Email address: wim.pouw@donders.ru.nl/babajide.owoyele@hpi.de

Website: <https://wimpouw.com>

Course info

Level: Introductory

Target group: 1st, 2nd Research Master's students, and PhDs in cognitive science and Linguistics. Participants have a broad general knowledge of linguistics. Participants with programming experience will benefit more from the course, but extensive programming experience is not needed for the conceptual level understanding of the course content.

Course description:

Wim Pouw and Babajide Owoyele will provide a deep dive into the latest developments in multimodal research. The session focuses on interaction between gesture-speech production, and embodied cognition. After a conceptual overview, we go into open science practices and computational tools, which includes using software for masking audiovisual data and extracting body and face motion features ([Owoyele, 2022, 2024](#)) with hands-on automatic analyses of multimodal signals.

Participants will engage in exercises drawing on [EnvisionBOX](#), a platform designed to facilitate multimodal analysis for early-career researchers who are beginning to program. These exercises will be about automated analysis pipelines that go through collection, automatic analyses, as well as archiving of data, such that your research is computationally reproducible as “open as possible, while as closed as necessary”. The session will equip students of communication with tools and methods that will be generally applicable in research projects in cognitive sciences, linguistics, also with bridging of qualitative and quantitative methods.

Day-to-day programme (ideally with an indication of hours to spend on each part)

Preparatory course work, to be done before Monday (22 hours in total)

- Preparatory reading material for the small preparatory assignment (e.g. 13 hours / 52 pages):
 - Pouw, W., Trujillo, J. P., & Dixon, J. A. (2020). The quantification of gesture–speech synchrony: A tutorial and validation of multimodal data acquisition using device-based and video-based motion tracking. *Behavior research methods*, 52, 723- 723-740. Doi (open access: [10.3758/s13428-019-01271-9](https://doi.org/10.3758/s13428-019-01271-9))
 - Pouw, W., Dingemanse, M., Motamedi, Y., Özyürek, A. (2021). A systematic

investigation of gesture kinematics in evolving manual languages in the lab.
Cognitive Science. doi (open access): [10.1111/cogs.13014](https://doi.org/10.1111/cogs.13014)

- Gregori, A., et al. (2023). A roadmap for technological innovation in multimodal communication research. HCII 2023.
https://www.researchgate.net/publication/372441859_A_Roadmap_for_Technological_Innovation_in_Multimodal_Communication_Research

- What do you want to learn to apply, and why?:
 - Assignment: Explore all modules on EnvisionBOX. Write down in a paragraph (max. half page) what modules on EnvisionBOX you find interesting and would like to learn, and how you would incorporate the module's outputs in a research idea you have. Send it to the instructors.

Course work on Monday (4 hours)

- Attend Monday's class (2 hours): *Quick primer about the computational multimodal analyses. Overview of open science, computational reproducibility, and ethical data Archiving.*
- Preparatory work: Getting started with python and jupyter notebook. We invite you to get started with running a jupyter notebook on your laptop, so you can run python code from EnvisionBox (feel free to use other software too, like Visual Studio Code). Here is a get started: <https://envisionbox.org/gettingstarted.html>. Use google and youtube to figure out any bugs that you run into along the way (e.g., "how to install python and Jupyter Notebook on OS"). Feel free to reach out to Wim Pouw or Babajide Owoyele if you are having any issues.

Course work on Tuesday (4 hours)

Attend Tuesday's class (2 hours): EnvisionBOX: Using Multimodal Data for Research

- First assignment (2 hours including readings, hand in before Wednesday's class):
 - Try to run the following script https://envisionbox.org/embedded_Mediapiping.html with a 5 second video you have downloaded or recorded with a speaking person in the frame. Send the video to the instructor to show you have motion tracking running. Google any package installation bugs that you find on the way, and try to solve your way out of it.

Course work on Wednesday (4 hours)

- Attend Wednesday's class (2 hours): EnvisionBOX: Using Multimodal Data for Research
- Second in-course assignment (2 hours including reading, hand in before Thursday's class):
 - Try to run the following script on your own video and try to summarize what it does as if writing it in a methods section of a scientific paper (half page)
https://envisionbox.org/embedded_AnimatingSoundMovement.html
(show the output video at class, and send the report to the instructor)

Course work on Thursday (4 hours)

- Attend Thursday's class (2 hours): Advanced Techniques in EnvisionBOX: Data Wrangling, Automatic Classification and Analyses
- Reading material in preparation for Friday's class (2 hours / 8 to 10 pages):
 - o Sievers, B., & Thornton, M. A. (2024). Deep social neuroscience: The promise and peril of using artificial neural networks to study the social brain. *Social Cognitive and Affective Neuroscience*, 19(1), nsae014.

Course work on Friday (4 hours)

- Attend Friday's class (2 hours): Discuss collaborative projects and knowledge sharing within the EnvisionBOX platform, focusing on expanding its capabilities.

Course work after Friday (40 hours)

- Final assignment (parts should add up to 40 hours):
 - Let's be creative and do some science fiction. We have overviewed relatively new methods in multimodal communication and behavior research (see www.envisionbox.org to review some methods, but feel free to be inspired by other methods covered in other courses as well). These methods still have limitations, but new breakthroughs are being made all the time at the moment. We want you to write a research proposal that makes use of methods that you think will be available in the next 10 years. So write a research proposal for the near future. First argue why you think method X or Y, or a combination of methods X and Y, will be available in the future, based on actual methods now out (max 1 page). And then proceed and write your research proposal of the future, with an introduction (with real references, and 10% fake references MARKED IN RED that are about the recent breakthroughs), and a methods section, with operationalizations of your concepts. Try to connect it to your theoretical interests.

What is an example of near-possible technology? Well we have overviewed advances in motion tracking from videos, or classifying behaviors like gestures from motion tracking, and perhaps you envisage improvements in these methods that allow you to do a unique experiment that you think is now not possible quite yet.

Please try to remain grounded in reality (no anti-gravity boots!).

Note sure about your idea? Feel free to contact the instructors or members of the class. We will also go over some examples in the course itself, to get a better idea of what we think is possible in the near future.

The final assignment will be sent to the instructor by the students themselves, somewhere between 4 to 8 weeks after the end of the course (the different LOT institutes have different RMI deadlines).

Further readings (optional):

- Maass, W., et al. (2018). Data-Driven Meets Theory-Driven Research in the Era of Big Data. <https://aisel.aisnet.org/jais/vol19/iss12/1/>
 - Some bigger picture thinking about how to think about types of data.
- Pouw, W., & Fuchs, S. (2022). Origins of vocal-entangled gesture. *Neuroscience & Biobehavioral Reviews*, 141, 104836.
<https://www.sciencedirect.com/science/article/pii/S0149763422003256>
 - Some in depth analysis on how movement science can provides new insights in classic topics in linguistics.
- Trujillo, J. P., Vaitonyte, J., Simanova, I., & Özyürek, A. (2019). Toward the markerless and automatic analysis of kinematic features: A toolkit for gesture and movement research. *Behavior Research Methods*, 51, 769-777.
<https://link.springer.com/article/10.3758/s13428-018-1086-8>
 - A paper that translates concepts of gesture studies into computationally reproducible measurements using motion tracking
- Alviar, Camila, Christopher T. Kello, and Rick Dale. "Multimodal coordination and pragmatic modes in conversation." *Language Sciences* 97 (2023): 101524.
<https://www.sciencedirect.com/science/article/pii/S038800012200064X>
 - A paper that relates the rise of new methods to possibilities for the complex systems approach to language.